The Race to Build Supermassive Black Holes

C. Tyler, B. Janus, D. Santos-Noble

Department of Physics, Fort Lewis College, Durango, CO 81301

tyler_c@fortlewis.edu

ABSTRACT

The high redshifts of the most distant known quasars, and the best estimates of their black hole masses, require that supermassive black holes (SMBHs) must have formed very early in history. Several mechanisms for creating and growing these holes have been proposed. Here we present an evaluation of the timescales needed for various critical processes in order to discriminate between the proposed scenarios. We find in particular that mergers alone are not able to grow the black holes at a sufficient rate. Accretion models offer a solution and we use accretion timescales to constrain the manner in which the black hole was first formed. This analysis implies, but does not require, the action of some unconventional process.

Subject headings: cosmology: theory – black holes – galaxies: evolution – quasars: general

1. Introduction

Central supermassive black holes are a common feature to galaxies today, but which came first, the black hole or the galaxy? Conventional thinking would suggest that the the first generation of stars evolved into black holes, which have subsequently settled to the centers of their host galaxies, merged, and accreted gas. But this idea, in which central black holes form inside pre-existing galaxies, has recently earned some scrutiny. First, the discovery of increasingly high redshift quasars requires a surprisingly early formation of the black holes (see, for example, Fan et al. (2001) and Fan et al. (2003)). Second, a large quasar sample shows no evidence of black holes growing in mass with decreasing redshift (Vestergaard 2002; Dietrich et al. 2002). So we are left to consider the possibility that either the central black holes formed before their host galaxies, or they grew to maturity very quickly within them. Either way, they have grown little since the quasar epoch.

The most distant known quasar lies at $z = 6.41$, with a central black hole of mass $M_\bullet = 3 \times 10^9$ $\mathrm{M}_\odot$ (Willott, McLure, & Jarvis 2003). In the ΛCDM cosmology observed by WMAP (Bennett et al. 2003), with $\Omega_\Lambda = 0.73$, $\Omega_\mathrm{m} = 0.27$, and $H_0 = 71$ km s^{-1} Mpc^{-1}, this redshift cor-

responds to a time when the universe was only 880 Myr old. For the present work, we will take this as the time to beat: 3 billion solar masses in 880 million years.

In the past year, two separate HST studies have cited evidence for intermediate mass black holes (IMBHs) in the centers of globular clusters: a 4000 $M_\odot$ hole in M15 (van der Marel et al. 2002), and a 20 000 $M_\odot$ hole in Andromeda's G1 cluster (Gebhardt, Rich, & Ho 2002). This is the lastest and strongest evidence for IMBHs, but there is additional evidence, and good theoretical motivation as well; see van der Marel (2003) for a comprehensive review. IMBHs are widely believed to be a necessary growth stage for SMBHs. In section 2 of this paper, we will review the major proposed routes to the formation of a SMBH, all of which include an IMBH phase, from which time the question is simply one of growth.

We start in Section 2 with a flowchart of avenues leading to the creation of a SMBH. In Section 3, we examine the timescales for each needed process. We conclude in Section 4 by discussing how realistic each avenue is in light of these timescales.

2. Flowchart

There are essentially four proposed families of models leading to the formation of IMBHs, and two or three ways to grow them. These approaches are depicted in figure 1 and discussed in turn below.

1. The black holes may be *primordial*, in which case they formed from primordial density variations before big bang nucleosynthesis.

Primordial black holes (PBHs) have been studied extensively, and the most direct mechanism for their creation is the collapse of gaussian random density fluctuations (Carr & Hawking 1974). These holes come from horizon scale (or smaller) modes, and therefore their masses are determined by their time of formation. In the radiation dominated early universe,

$$M_\bullet \simeq 10^5 \left(\frac{t}{\mathrm{s}}\right) \ \mathrm{M}_\odot \ . \qquad (1)$$

But in order to preserve the successful BBN prediction of light element abundances, there must be no significant rate of PBH formation once nucleosynthesis begins, and therefore the PBHs are capped at intermediate mass. In addition, Carr, Gilbert, & Lidsey (1994) have pointed out that, given a small scalar spectral index – $n_s = 0.93 \pm 0.03$ was recently observed in the CMB (Bennett et al. 2003) – PBHs from density inhomogeneities should only have formed in quantities too small to be of interest.

A more promising, and perhaps inevitable mechanism for forming PBHs also exists, in which the collapse is triggered by "color holes" at the quark-hadron phase transition (Crawford & Schramm 1982). However, because this occurred at $\sim 10^{-6}$ s, these PBHs would be smaller than $\sim 1\ M_\odot$ by eq. 1, and would remain as collisionless dark matter today, rather than collecting into larger black holes. (Interestingly, Hawkins (1996) shows evidence for such PBHs in the microlensing of distant quasars, in numbers comparable to that needed to account for dark matter.)

2. Normal *population III stars* formed at $\sim 100\ M_\odot$, evolved to black holes, and merged at the center of their small dark matter halos.

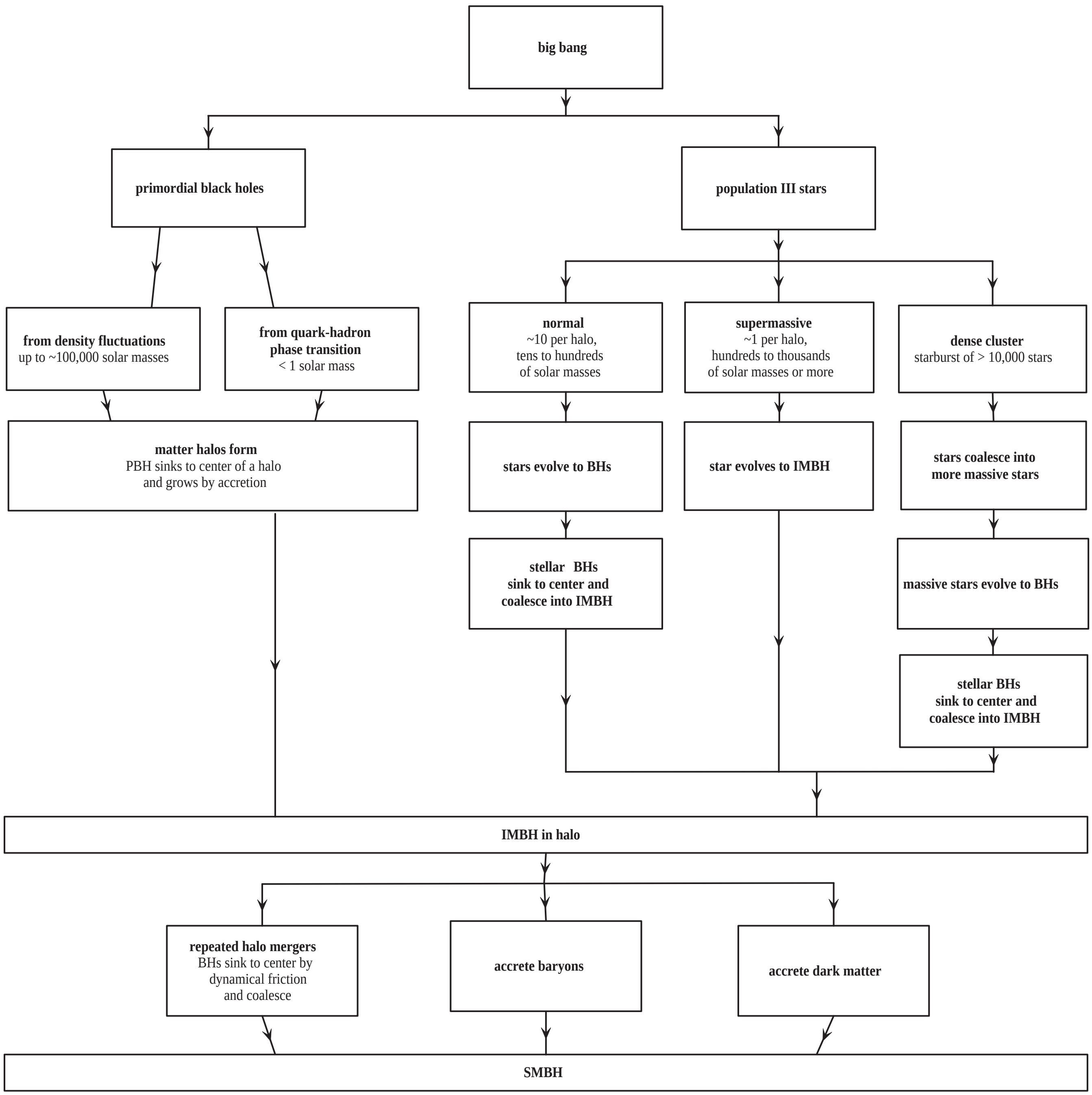

Fig. 1.— The various ways to build the earliest supermassive black holes in the universe.

This is perhaps the least exotic way to create IMBHs, and at this point there is very strong support for the process in the form of numerical simulations of structure formation (Abel, Bryan, & Norman 2000; Yoshida et al. 2003). These simulations include the relevant atomic and molecular processes in the first gas clouds, particularly cooling by rotation of molecular hydrogen, superimposed on CDM halo evolution. They find that $\lesssim 10$ stars of ~ 100 $M_\odot$ form in halos of $\sim 10^6$ $M_\odot$, engaging $\lesssim 1\%$ of the system's baryonic matter.

Because the cooling of population III stars is hindered by the lack of metals, these first stars would be more massive than those allowed by fragmented star formation today. Heger & Woosley (2002) have shown that these massive stars will evolve into black holes containing a significant fraction of the star's initial mass (this fraction varies with the stellar mass, and is of order 50%), unless the stars are in the mass range 140 $M_\odot < M <$ 260 $M_\odot$, in which case they are completely disrupted when they go supernova. Given the small initial size of such a cosmologically young halo (~ 10 pc), the holes sink to the center and presumably merge into a single IMBH there.

3. *Supermassive stars* may have been the first baryonic objects to form. If so, they would have evolved rapidly into IMBHs.

Supermassive stars (SMSs), as a class of objects, span from 10^3 to 10^8 $M_\odot$, although the first generation of them would reside at the lower end of that range (Shapiro & Teukolsky (1983) provides a comprehensive introduction on the topic). A 10^3 $M_\odot$ SMS has a lifetime of $\lesssim 10^6$ years (SMS lifetimes range down to 10 years), at the end of which, it undergoes relativistic collapse to a black hole in a matter of seconds (New & Shapiro 2001). This collapse sends $\sim 90\%$ of the star's original mass into the black hole, with the remaining 10% forming an accretion disk (Shapiro & Shibata 2002; Shibata & Shapiro 2002).

The radius of a 10^3 $M_\odot$ SMS, $R_{\rm SMS} \approx 0.1$ pc, is around 10% of the virial radius for the earliest halos containing gas cool enough to collapse. It is therefore reasonable to suspect that these stars should appear in early halos, especially since population III stars produce ultraviolet radation which splits hydrogen molecules, thus destroying the only coolant which could collapse smaller stars. The collapsing gas might produce enough radiation pressure to prevent fragmentation (Baumgarte & Shapiro 1999). Another possibility is that the first stars formed dense clusters which merged into one or more SMSs. However, whether or not supermassive stars ended the cosmological dark ages remains a matter of speculation.

4. If *dense clusters* of stars emerged from early star forming regions, then the cluster stars may have merged with each other to create massive stars (say, several hundred solar masses), which then evolved to massive seed black holes. This idea has been mapped out recently by Ebisuzaki et al. (2001), and we will not repeat it here. However, for the sake of timescales, which are the focus of this paper, this scenario is nearly the same as the population III star scenario discussed above, because the merging occurs during the lifetime of the cluster's massive stars.

Each of the paths in figure 1 sees a phase with an IMBH imbedded in a halo, at which point accretion and halo mergers take over to achieve heavier black holes.

The merger scenario merges central black holes hierarchically along with the host halos, growing both together. It is clear that merging IMBHs and SMBHs are real, interesting phenomena deserving study. For example, Milosavljević & Merritt (2001) have simulated the whole process, and shown that coalescing SMBHs remove the steep cusps that CDM simulations generate, thus matching observed galaxy rotation curves. But it remained to be shown here whether or not mergers are prevalent enough to sufficiently increase black hole mass.

Gas accretion is another major way to grow IMBHs into SMBHs. The belief that quasars are powered by accretion at the Eddington luminosity gives this idea credibility. But over what fraction of cosmic history did these holes accrete? Some authors suppose that Eddington limited accretion is continuous since the black hole first formed (Haiman & Loeb 2001), and thus its mass grows rapidly. Others have proposed scenarios in which accretion is negligible over the history of the BH (Madau & Rees 2001), such that 150 $M_\odot$ holes that formed in the first episode of star formation still populate our Galactic bulge today.

Another possibility is that the BH feeds on dark matter, in which case the halo and the IMBH evolve together. MacMillan & Henriksen (2002) proposed a model in which the halo and hole grow self-similarly, with both analytical and numerical results (MacMillan & Henriksen 2003) supporting their idea. This model is particularly interesting because it also simultaneously explains both the spectral index for the galaxy power spectrum, $n \approx -2$, and the $M_\bullet \propto \sigma^4_{\rm bulge}$ relationship observed to be universal in galaxy centers (Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002).

All of these methods certainly contribute to the growth of massive black holes. In the next section we will use timescales to help constrain which methods could realistically dominate the process of forming SMBHs in less than a gigayear.

3. Timescales

First, we examine the evolution to the IMBH phase, as depicted in Figure 1. If the BH is primordial, then it can be up to 10^5 $M_\odot$ in the very early universe. Its growth by merger or accretion then awaits the formation of matter clumps to surround the PBH. These could begin to arrive as early as photon decoupling, so timescales alone will not strongly constrain this possibility unless the PBH in question started at < 1 $M_\odot$, as it would if it formed during the quark-hadron phase transition. In this case the BH seed has a very long way to grow, which is barely possible by accretion, as will be shown in subsection 3.3.

The simulations of population III star formation of Yoshida et al. (2003) produce some very useful information. The earliest simulated star forming region occurs at $z = 32$, with stars forming at 10^{-3} $M_\odot\,{\rm yr}^{-1}$ per comoving ${\rm Mpc}^3$, increasing to 0.1 $M_\odot\,{\rm yr}^{-1}\,{\rm Mpc}^{-3}$ at $z = 7$, and holding constant thereafter (Hernquist &

Springel 2003).

Molecular hydrogen coolant forms in these clouds in $\sim$ 30 Myrs. By $z = 23$, the gas pressure has diminished and gas clumps begin to contract. Halos are $\sim 5 \times 10^5$ to 2×10^6 $M_\odot$ by this time, and the largest virialized star forming clouds are 10^5 $M_\odot$. The gas takes an approximately isothermal density profile, with constant density inside a core radius of $\sim$ 10 pc.

The virial temperature is a few thousand Kelvins. A simple application of the virial theorem gives a virial radius of 45 pc, and the corresponding circular orbital speed is 10 km s^{-1}.

This, then, is the starting point for population III stars leading to IMBHs, with star formation beginning in earnest at $z = 23$. To convert redshifts to times in a ΛCDM cosmology with $\Omega_0 = 1$ (Peacock 1999),

$$t_{\rm age} = \frac{2}{3} H_0^{-1} \frac{\sinh^{-1} \sqrt{|\Omega_{\rm m} - 1|(1+z)^{-3}\Omega_{\rm m}^{-1}}}{\sqrt{|\Omega_{\rm m} - 1|}} \tag{2}$$

and with WMAP values inserted, this is

$$t_{\rm age} = 10.8 \sinh^{-1} \left(\sqrt{2.7\,(1+z)^{-3}} \right) \ {\rm Gyr}\,. \tag{3}$$

So gas collapses into the first stars at $t_{\rm age} \approx$ 0.15 Gyr.

These stars collapse in the free fall time,

$$t_{\rm ff} = \sqrt{\frac{3\pi}{32\,G\rho}}\,, \tag{4}$$

which is about 30 Myr at the densities observed in simulations by Abel, Bryan, & Norman (2000). They evolve to black holes (or disruptive supernovae) in a few million years, leaving $\sim$ 0.7 Gyr remaining to grow into a quasar power source.

Using the dynamical friction formula from Chandrasekhar (1943), quoted here from Binney & Tremaine (1987), we can find the time needed for these population III remnant BHs to sink to the center of their host halo. Treating the BHs as point particles,

$$t_{\rm df} = \frac{1.17}{G \ln \Lambda} \frac{v r_0^2}{M_\bullet}\,, \tag{5}$$

where $\ln \Lambda$ is called the Coulomb logarithm, which characterizes the distance range perpendicular to the path of motion over which the frictional encounters occur (for a recent study of this parameter, see Spinnato, Fellhauer, & Portegies Zwart (2003)). Using the values mentioned above for the parameters in this equation, and choosing $\Lambda \approx 20$ (this choice is common and doesn't affect the outcome much because of the logarithm), the dynamical friction timescale relevant here is at least $\sim$ 0.1 Gyr, bringing the total down to $\sim$ 0.6 Gyr available for BH growth. The coalescence time of these holes is not well known, and could be long.

If the star formation yields one or more supermassive stars, rather than multiple normal population III stars, then the evolution to a BH is faster than for normal stars (Shapiro & Teukolsky 1983). The dynamical friction time to merge smaller stellar BHs is also avoided with SMSs, so in the SMS case, we have $\sim$ 0.7 Gyr remaining to grow a SMBH.

In the case of an early-forming dense star cluster in which runaway collisions merge the stars, the timing remains essentially the same. One waits until the cluster stars form, merge, and evolve to one or more BHs. The time available to become supermassive is still at most 0.7 Gyr.

3.1. Dynamical Friction

We consider the case of a satellite subhalo halo falling into a larger halo. Each halo is modeled as a singular isothermal sphere, so that $\rho \propto r^{-2}$ and orbital velocities within a halo are constant. We seek a formulation that tracks the friction until the satellite's central BH reaches the new center.

Smaller halos formed earlier in the CDM hierarchy, when the universe was denser, and therefore they are more tightly bound. So as a subhalo sinks, we assume it holds itself together, until it gets close to the center of the larger halo. There, which halo an orbiting particle belongs to becomes ambiguous. This occurs roughly when a particle's distance from its original host subhalo is on order of the separation between the two halo centers. This separation distance is r, a radial coordinate extending from the center of the larger halo.

The internal speed within an isothermal halo is

$$v = \sqrt{\frac{GM}{R}}\,. \tag{6}$$

The square root in this equation is helpful because it reduces the dependence that velocities have on a halo size R, which is generally not well known.

The mass which can be associated with the sinking satellite is then

$$M_{\rm s} \simeq M_r + M_\bullet = \frac{v_{\rm s}^2 r}{G} + M_\bullet\,, \tag{7}$$

where M_r is the mass of the satellite halo residing inside a distance equal to r around it, $M_\bullet$ is the satellite subhalo's central BH mass, and $v_{\rm s}$ is the orbital velocity inside the satellite. Choosing to use r in this way assumes that the isothermal satellite halo extends to distances larger than some initial infall radius: $R_{\rm s} \geq r_i$.

The angular momentum of the satellite in the larger halo is $\vec{L} = \vec{r} \times M_{\rm s}\vec{v}$, where $\vec{v}$ is the orbital velocity within the larger halo. Dynamical friction induces a torque $\vec{\tau} = \vec{r} \times \vec{F} = \vec{r} \times M_{\rm s}(d\vec{v}/dt)$.

Following the standard derivation presented in Binney & Tremaine (1987), the statement of $\vec{\tau} = d\vec{L}/dt$ leads to the equation

$$r\frac{dr}{dt} = -0.428\,\frac{GM_{\rm s}}{v}\ln\Lambda\,. \tag{8}$$

We next define 3 constants as follows: $C_1 = (0.428/v)\ln\Lambda$, $C_2 = C_1 G M_\bullet$, and $C_3 = -C_1 v_{\rm s}^2$. Inserting the satellite mass from eq. 7, we have

$$\frac{dr}{dt} + C_2 r^{-1} = C_3\,. \tag{9}$$

Eq. 9 is a first-order, linear, ordinary differential equation, but the closed-form solution is prohibitively long and complicated due to the r^{-1} factor in the second term. A numerical solution readily parameterizes the sinking orbit $r(t)$, and a root can be found for the time $t_{\rm df}$ at which $r \to 0$. One needs only to specify a starting radius for the satellite halo, $r_i \equiv r(0)$. An example orbit decay curve computed in this way is given in figure 2.

This dynamical friction model approximately doubles the time spent spiralling inward, relative to the Chandrasekhar formula for an infalling point particle, eq. 5. Colpi, Mayer, & Governato (1999) found the same result using both N-body simulations and the theory of linear response. A subsequent paper by Taffoni et al. (2003) gives an analytical approximation for NFW halos (Navarro, Frenk, &

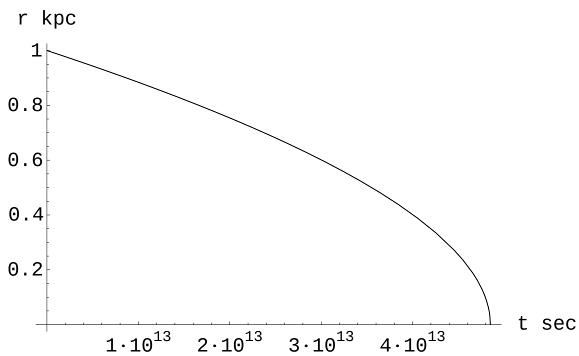

Fig. 2.— Dynamical friction infall profile as calculated in section 3.1. In this figure we have exaggerated the curve at the end for clarity by using an unrealistically large BH mass in the satellite subhalo.

White 1996), also finding that the decay time is increased by a factor of a few if the satellite has less than $\sim 10\%$ of the larger halo's mass. In the mass range of interest for the present work, all these models essentially agree; however, since we wish to follow length and mass scales ranging down to a central BH, we have chosen to use the method described above.

3.2. Mergers

We first consider the viability of black hole merging as a means to grow SMBHs. As halos merge over cosmic history, their central BHs sink to the new center by dynamical friction, and subsequently coalesce into a larger BH.

The usual way to track mergers is by N-body simulation. We begin with such a simulation, and calculate the total mass of subhalos that can merge into one large halo between the epoch of the first massive BHs and the early quasar epoch. We subject this calculation to the constraint that the merging subhalos must sink to the center in the time allowed between the two epochs.

This can become a very complicated calculation, because the halo population is continuously changing over a long period of cosmic time. To make the problem tractable, we make a series of major approximations. We describe them along the way, and discuss them afterward. We demonstrate first that with our assumptions, mergers are *not* sufficient to grow SMBHs, regardless of how the BHs originally formed. Second, we show that a more accurate calculation to replace our simplifying assumptions would only strengthen that conclusion.

The argument begins with the numerical simulations of Volonteri, Haardt, & Madau (2003). They follow a 4×10^{13} $M_\odot$ halo today back in time to all its subhalos, presenting halo mass spectra at multiple redshifts. For our purposes, the results they present at $z = 5$ and $z = 20$ are most important, approximating the time of the first quasars ($z = 6.41$) and the first IMBHs in halos (after $z = 23$). We'll address the impact of these approximations in redshift and total mass after the calculation.

The halo functions are nearly power law in the mass range of interest. For M in $M_\odot$ units, at $z = 5$

$$\frac{dN}{dM} = 10^{14}\, M^{-\frac{7}{3}}\,, \tag{10}$$

and at $z = 20$

$$\frac{dN}{dM} = 1.6 \times 10^{16}\, M^{-2.9}\,. \tag{11}$$

Consider the largest halo in existence as the period of interest begins. Imagine (conservatively) that all other subhalos are

available to merge into this one, and (also conservatively) that all of these subhalos contain central BHs of 0.1% of the parent galaxy's luminous mass, as observed in large galaxies today (Ferrarese & Merritt 2002). According to WMAP, the ratio of baryonic matter to total matter is 0.17 (Bennett et al. 2003), so conservatively we may say that a central BH has 0.02% of the host subhalo's mass. Then every subhalo massive enough to dynamically sink to the center in the available time does sink, leading to a more massive combined halo.

If all the IMBHs from all the combined subhalos coalesce quickly (again, highly conservative), then the new SMBH mass is 0.02% of the total mass of all subhalos with mass $M \gtrsim M_{\rm s,min}$, the minimum satellite mass needed for dynamical friction to sink a subhalo by the early quasar era.

However, during the merging process, the largest halo is gaining mass and the number of massive halos is increasing. To evade this problem, we suppose that the halo mass spectrum at the *end* of the merging process (at $z = 5$) applies throughout the merging process. Our calculations show that in most cases, this assumtion conservatively exaggerates the final SMBH mass by more than an order of magnitude.

Using eq. 10, the total mass of the quasar's SMBH is

$$M_\bullet = 0.0002 \int_{M_{\rm s,min}}^{\infty} 10^{14}\, M^{-\frac{4}{3}}\, dM\,. \quad (12)$$

The lower limit, $M_{\rm s,min}$, is determined by dynamical friction as described in section 3.1. We follow the infalling halo mass, and transition to the (smaller) infalling BH mass as the two holes close in, achieving results closely consistent with other models that include mass loss but not a central BH (Colpi, Mayer, & Governato 1999).

In the dynamical friction calculation, three parameters must be chosen. We take the internal velocity of the larger halo to be $v \approx 100$ km s^{-1}, noting that for a fixed dynamical sinking time (fixed by the time available for BH growth), the final SMBH mass obeys $M_\bullet \propto v^{-1/3}$. So choosing a larger velocity would only very slightly reduce the final SMBH mass by reducing the number of subhalos which sink in the allowed time.

The other needed quantities are radii: r_i, the initial radius from which the satellite halo begins its descent, and $R_{\rm s}$, the size of the satellite subhalo. The former is roughly the impact parameter of the colliding subhalos, and the latter indicates the subhalo's concentration factor. Both radii will clearly vary from one encounter to the next, so we choose several representative values. The results demonstrate that any reasonable (in fact, conservative) choice of these numbers leads to the same conclusion, that merging black holes take too long.

As shown earlier, the timescale between the first IMBH and a quasar's SMBH depends on how the BH formed. A primordial BH has the age of the universe at the time of the most distant quasar's emission known today, or 0.88 Gyr. Population III stars and supermassive stars have 0.7 Gyr. If the population III stars merged within their original halo first, then they have 0.6 Gyr to grow.

We fix the sinking time at the appropriate value and compute the total BH mass that can be assembled by halo mergers, neglecting the time needed for coalescence.

t (Gyr)	R_s (kpc)	r_i (kpc)	$M_\bullet$ ($M_\odot$)
0.88	10	10	4.8×10^7
	10	1	1.0×10^8
(PBH)	10	0.1	2.2×10^8
	1	1	2.2×10^8
	1	0.1	4.8×10^8
0.7	10	10	4.4×10^7
	10	1	9.5×10^7
(pop III,	10	0.1	2.1×10^8
cluster,	1	1	2.1×10^8
or SMS)	1	0.1	4.4×10^8
0.6	10	10	4.2×10^7
	10	1	9.0×10^7
(merged	10	0.1	2.0×10^8
pop III)	1	1	1.9×10^8
	1	0.1	4.2×10^8

Table 1: SMBH masses achieved through halo mergers by the early quasar epoch. R_s is a characteristic size for the satellite subhalos, and r_i is the radius from which they are assumed to begin their infall. t is the time allowed for growth.

Our results are listed in Table 1, which lists a maximum SMBH mass for various choices of R_s and r_i. No choice brings $M_\bullet$ up near 3×10^9 $M_\odot$, the value needed to explain the most distant known quasars.

One is led to conclude that mergers alone are not sufficient to grow the largest SMBHs in the oldest quasars. Note that although the argument above relies on many simplifying assumptions, all but one tend to overestimate the resulting SMBH mass, so the SMBH masses in table 1 should be considered extreme upper limits. Our assumptions are summarized as follows.

1. All halos contain central BHs at 0.02% of the halo mass (this may seriously overestimate the number of BHs).

2. All relevant merging is completed by $z = 5$, instead of today, so the number of infalling BHs is again exaggerated.

3. In the ΛCDM world, going from redshift 20 to 5 allows over 1 Gyr for merging, so the number of more massive subhalos available to merge is artificially large. This overestimates the final SMBH mass, because larger infalling subhalos will dynamically sink faster.

4. Infall radii begin at $r_i \leq R_s$, which for some mergers represents a head start of multiple orders of magnitude.

5. The choice of an isothermal halo model probably exaggerates the central density, artificially speeding up the action of dynamical friction near the center.

6. SMBH coalescence is assumed to occur arbitrarily quickly, but in reality takes at least $\sim 10^8$ years, and probably much longer (Milosavljević & Merritt 2001).

7. The choice to trace a 4×10^{13} $M_\odot$ halo today back in time tends to *under*estimate the number of progenitor subhalos. Extrapolating this mass linearly from Volonteri, Haardt, & Madau (2003), we find that increasing it by 100 times would increase dN/dM at $z = 5$ by a factor of $\lesssim 10$ while approximately maintaining its slope.

In this case, the largest halo at $z = 5$ would be $\sim 10^{13}$ $M_\odot$, so halo speeds should be comparable to those in M87 today, the giant elliptical at the center of the Virgo galaxy cluster. Radial velocities for globular clusters around M87 (Cohen 2000) range from 1000 to 2000 km s^{-1}. These changes, using $v \approx 1000$ km s^{-1}, allow a SMBH to become almost 10^9 $M_\odot$ in 0.7 Gyr from an initial radius $r_i = R_s =$

1 kpc, which is a factor of ~ 4 greater than our original result.

So if we attempt to correct for this seventh assumption, we approach (but do not reach) the 3×10^9 $M_\odot$ hole that quasar observations require. But this is only true if we take literally the six highly conservative assumptions above.

3.3. Accretion

Consider first gas accretion, followed by feeding on dark matter.

Gas accretion is probably the simplest and most successful way to explain the growth of an IMBH into a SMBH, or even a large stellar mass BH into a SMBH. Haiman & Loeb (2001) show that if the black hole accretes at the Eddington limit continuously, from birth through detection as a quasar, it achieves the needed growth. They give the Eddington luminosity,

$$L_{\rm E} = 4\pi G M_\bullet c \mu_{\rm e} m_{\rm p} \sigma_{\rm T}^{-1}\,, \qquad (13)$$

and introduce a radiative efficiency $\epsilon \equiv L/\dot{M}_\bullet c^2$ and a fraction of Eddington output $\eta \equiv L/L_{\rm E}$. The resulting e-folding time is

$$t_{\rm acc,e} = \frac{M_\bullet}{\dot{M}_\bullet} = 4 \times 10^7 \left(\frac{\epsilon}{0.1}\right) \eta^{-1}\,{\rm yr}\,, \quad (14)$$

which equals a 10-folding timescale

$$t_{\rm acc,10} = 9.2 \times 10^7 \left(\frac{\epsilon}{0.1}\right) \eta^{-1}\,{\rm yr}\,. \qquad (15)$$

For a fiducial value of $\epsilon \approx 0.1$ and continuous Eddington-limited accretion ($\eta \approx 1$), one can calculate the initial BH mass needed to grow to 3×10^9 $M_\odot$ in the alloted time.

If the seed BH is primordial, it might accrete steadily from approximately the beginning of time. In this case, it has 0.88 Gyr to grow; that's 9.6 orders of magnitude in mass by eq. 15. So the initial seed mass must have been $M_{\rm seed} \gtrsim 0.8$ $M_\odot$, which, interestingly, could have been formed during the cosmological quark-hadron phase transition (Jedamzik & Niemeyer 1999).

If the seed BH is a normal population III stellar remnant, or a supermassive star's remnant, then it had 0.7 Gyr. It's initial mass was then $M_{\rm seed} \gtrsim 70$ $M_\odot$, which is completely plausible for the first generation of stars. One star of $\gtrsim 260$ $M_\odot$ would suffice (Heger & Woosley 2002).

In 0.6 Gyr, the seed BH mass needs to be $M_{\rm seed} \gtrsim 900$ $M_\odot$. This applies for the case where population III stars must sink to the center of their parent halos and merge before beginning to grow significantly by gas accretion.

In the case where the seed BH formed from the evolution of normal population III stars, the mass of the seed star is suggestively close to what one would expect. Thus this route is possible, but only if Eddington limited accretion is maintained during the entire IMBH growth process. If the seed hole evolved from a more massive SMS, then there is time to spare, and the accretion rate could have dropped for some of the growth period.

The success of this quick exponential growth calculation seems to indicate that baryon accretion is the most realistic mechanism for IMBH growth. When the BH is small, the Eddington mass accretion rate is not prohibitively large, and by $z = 6.41$ we have very strong evidence that the quasar is accreting at its Eddington rate (Willott, McLure, & Jarvis 2003). So it is reasonable

to speculate that it has been accreting at this rate all along.

However it is worth noting that the central SMBHs in modern galaxies appear to grow by a factor of 10 on timescales of $\gtrsim 8$ Gyr (Merrifield, Forbes, & Terlevich 2000). This is very slow compared to Eddington limited growth, requiring $\eta \lesssim 0.01$. This dramatic drop in accretion rate needs explaining, especially if we are to claim that no similar drop has ever happened before $z \sim 6$ in all of cosmic history, as the notion of SMBH growth by gas accretion onto population III stellar corpses would mandate. This problem is only alleviated if the first stars were supermassive, or the seed BHs were primordial and significantly larger than 1 $M_\odot$.

Accreting dark matter is another possibility. MacMillan & Henriksen (2002) have proposed an interesting way to account for IMBH growth in which the hole grows proportionally to the dark matter halo as matter falls in. As mentioned earlier in section 2, this assumption leads to the observed galaxy power spectrum and the observed $M_\bullet - \sigma$ relation.

In numerical simulations (MacMillan & Henriksen 2003), dark matter was found to form a self-similar region surrounding the BH, which does grow in proportion to the BH growth at the center. The formation of this region proceeds by gravitational interactions between clumps of dark matter ("particles" in the simulations) which must have formed earlier in the CDM universe. The BH grows whenever the particles cross inside the Schwarzschild radius. (The BH growth rate is thus understated in the sense that $R_{\rm sch}$ was used instead of the larger capture radius.)

In the absence of angular momentum, the black hole grows rapidly: $M_\bullet \propto t^4$. But for realistic departures from pure spherically symmetric infall, in which the dark matter has angular momentum to help it resist falling into the hole, the growth of the BH is only linear in time: $M_\bullet \propto t$ (although the $M_\bullet - \sigma$ relation is recovered either way).

The simulations were not performed in standard units, and therefore cannot be immediately applied to relevant halo sizes and cosmic timescales. Here we simply note that the simulated BH growth is slow (e.g., linear in time, as compared with exponential in time for gas accretion). One therefore needs an unusually large IMBH to start from, which requires an unusually large SMS as its origin.

4. Conclusions

In this paper we have surveyed the state of knowledge of SMBH formation and growth in light of a difficult new data point to satisfy: a quasar at $z = 6.41$ whose SMBH was 3×10^9 $M_\odot$ when the universe was only 880 Myr old. We have extended the calculations of others on SMBH growth via mergers and accretion.

In the case of SMBH growth by mergers, we follow the output from a CDM simulation in order to track progenitor subhalos. We stipulate that a subhalo must sink to the center by dynamical friction after a merger within the time available for SMBH growth. This limits the mass of subhalos and therefore BHs which could have coalesced in time, thus constraining the final SMBH mass. We find that in any realistic case, mergers are incapable of growing a BH with the needed speed.

Accreting baryonic matter represents a viable growth process. For this mechanism, we impose time limits sensitive to the manner in which the seed BH first formed. The following possibilities emerge:

1. A primordial black hole that formed during the quark-hadron phase transition has accreted at the Eddington rate continuously ever since baryons have been able to cluster until at least the quasar epoch.

2. A large population III seed star with $M \approx 260\ \mathrm{M}_\odot$ evolved to a BH, and accreted at the maximum (Eddington) rate continuously between its formation and the quasar epoch.

3. A collection of merged population III stellar black holes formed an IMBH with $M \approx 900\ \mathrm{M}_\odot$, and accreted at the Eddington rate continuously through the quasar epoch.

Each of these options involves seed stars at the upper end of their predicted mass range. So although the Eddington limited gas accretion scenario is adequate to grow the holes in the time available, it is not entirely satisfying, because it requires heavy gas infall which is never significantly diminished before the quasar era, despite the relative gas suffocation which has evidently happened since. One can avoid imposing this requirement by starting with a larger BH:

4. A single SMS, or the merged remains of a dense cluster of stars, or a PBH formed well after the quark-hadron phase transition, yielded an IMBH much larger than $\sim 100\ \mathrm{M}_\odot$. It accreted through the quasar epoch, but did *not* need to maintain the maximum accretion rate during that time.

One other option, wherein the BH grows in proportion to its host halo by eating dark matter, is expected to yield slower BH growth. As such, normal population III stellar seeds would be inadequate, although a particularly massive SMS seed remains a possibility.

We conclude by suggesting that *some* unconventional mechanism is needed to realistically beat the clock for the early formation of SMBHs. BH growth through mergers is too slow, and growth by accreting nonstop at the Eddington limit is probably too contrived. The remedy appears to require one of the following: (a) PBHs from unusually large overdensity pockets, (b) population III starbursts well in excess of those predicted by current simulations, or (c) supermassive stars.

This work was supported by NASA under the Colorado Space Grant Consortium. We are grateful to R. N. Henriksen and J. D. MacMillan for helpful correspondence.

REFERENCES

Abel, T., Bryan G. L., and Norman, M. L., Astrophys. J. 540 (2000) 39.

Baumgarte, T. W. and Shapiro, S. L., Astrophys. J. 526 (1999) 941.

Bennett, C. L. et al. (WMAP collaboration), Astrophys. J. in press, astro-ph/0302207.

Binney, J. and Tremaine, S. "Galactic Dynamics," Princeton: Princeton University Press, 1987.

Carr, B. J., Gilbert, J. H., and Lidsey, J. E., Phys. Rev. D 50 (1994) 4853.

Carr, B. J., and Hawking, S. W., MNRAS 168 (1974) 399.

Chandrasekhar, S., Astrophys. J. 97 (1943) 255.

Cohen, J., Astron. J. 119 (2000) 162.

Colpi, M., Mayer, L., and Governato, F., Astrophys. J. 525 (1999) 720.

Crawford, M. and Schramm, D., Nature 298 (1982) 538.

Dietrich, M., et al., Astrophys. J. 581 (2002) 912.

Ebisuzaki, T. et al., Astrophys. J. 562 (2001) L19.

Fan, X., Astron. J., 122 (2001) 2833.

Fan, X., Astron. J. in press, astro-ph/0301135.

Ferrarese, L., and Merritt, D., Astrophys. J. 539 (2000) L9.

Ferrarese, L., and Merritt, D., Phys. World 15N6 (2002) 41.

Gebhardt, K., Rich, R. M., and Ho, L. C., Astrophys. J. 578 (2002) L41.

Gebhardt, K., et al., Astrophys. J. 539 (2000) L13.

Haiman, Z. and Loeb, A., Astrophys. J. 552 (2001) 459.

Hawkins, M. R. S., MNRAS 278 (1996) 787.

Heger, A. and Woosley, S. E., Astrophys. J. 567 (2002) 532.

Hernquist, L. and Springel, V., MNRAS 341 (2003) 1253.

Jedamzik, K. and Niemeyer, J. C., Phys. Rev. D 59 (1999) 124014.

MacMillan, J. D. and Henriksen, R. N., Astrophys. J. 569 (2002) 83.

MacMillan, J. D. and Henriksen, R. N., "Carnegie Observatories Astrophysics Series, vol. 1: Coevolution of Black Holes and Galaxies," ed. Ho, L. C., 2003, in press, astro-ph/0212278.

Madau, P. and Rees, M. J., Astrophysy. J. 551 (2001) L27.

McLure, R. J. and Jarvis, M. J., MNRAS in press, astro-ph/0204473.

Merrifield, M. R., Forbes, D. A., and Terlevich, A. I., MNRAS 313 (2000) L29.

Milosavljević, M., and Merritt, D., Astrophys. J. 563 (2001) 34.

Navarro, J. F., Frenk, C. S., and White, S. D. M., Astrophys. J. 462 (1996) 563.

New, K. C. B. and Shapiro, S. L., Class. Quant. Grav. 18 (2001) 3965.

Peacock, J. A., "Cosmological Physics," Cambridge: Cambridge University Press, 1999.

Shapiro, S. and Teukolsky, S., "Black Holes, White Dwarfs, and Neutron Stars," New York: John Wiley and Sons, 1983.

Shapiro, S. L. and Shibata, M., Astrophys. J. 577 (2002) 904.

Shibata, M. and Shapiro, S. L., Astrophys. J. 572 (2002) L39.

Spinnato, P. F., Fellhauer, M., and Portegies Zwart, S. F., MNRAS 334 (2003) 22.

Taffoni, G., Mayer, L., Colpi, M., and Governato, F., MNRAS 341 (2003) 434.

Tremaine, S. et al., Astrophys. J. 574 (2002) 740.

van der Marel, R. P., "Carnegie Observatories Astrophysics Series, vol. 1: Coevolution of Black Holes and Galaxies," ed. Ho, L. C., 2003, in press, astro-ph/0302101.

van der Marel, R. P., Gerssen, J., Guhathakurta, P., Peterson, R. C., and Gebhardt, K., astro-ph/0209314.

Vestergaard, M., Astrophys. J. 571 (2002) 733.

Volonteri, M., Haardt, F., and Madau, P., Astrophys. J. 582 (2003) 559.

Willott, C. J., McLure, R. J., and Jarvis, M. J., Astrophys. J. 587 (2003) L15.

Yoshida, N., Abel, T., Hernquist, L., and Sugiyama, N., Astrophys. J. in press, astro-ph/0301645.

This 2-column preprint was prepared with the AAS LaTeX macros v5.0.